

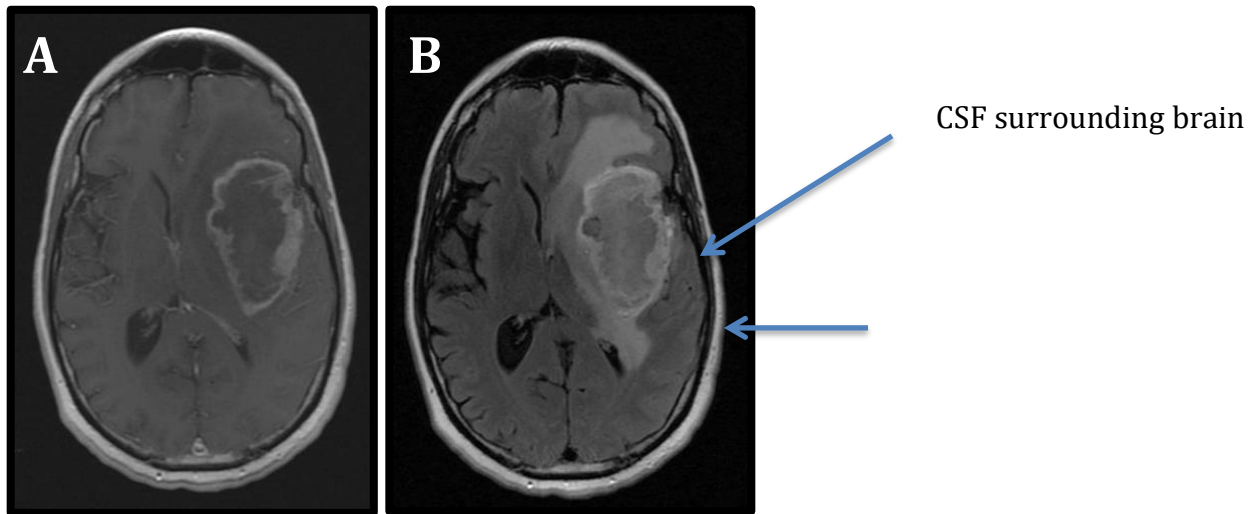
Life Sciences Outreach Faculty Speaker Series for High School Biology Teachers
How Biologists View Structure and Function
Fall 2018

Station 1: Imaging informs structure and function for doctors

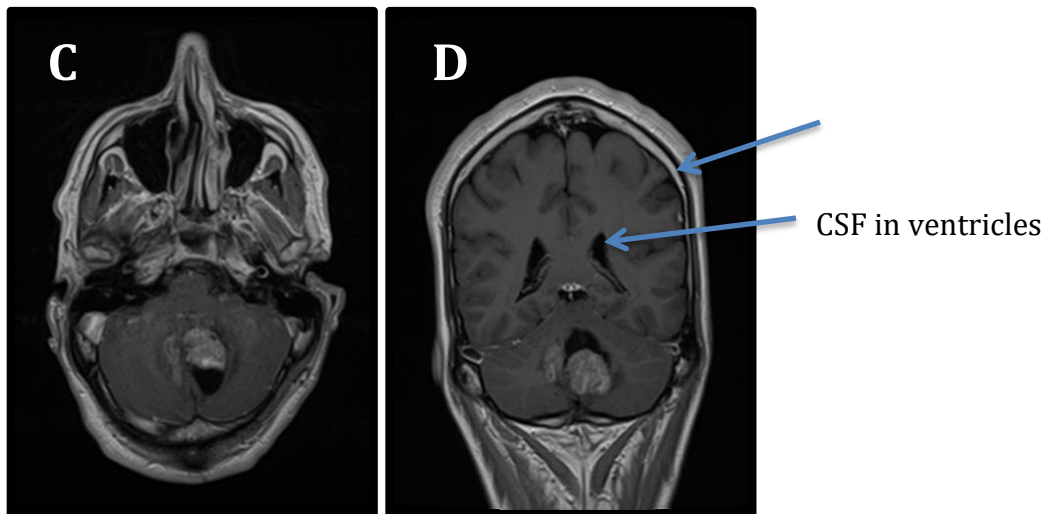
Imagine you are a doctor, and two patients come to you complaining of different symptoms. You order MRIs and get the following images. Your goal is to look at the MRIs and describe them using anatomical terminology. Finally, examine the brain sections, and match the brain sections to the MRI to correctly label the type of tumor present in the brain.

1. Examine the MRIs below and answer the questions:

Patient #1: MRI A & B Symptoms: confusion, sleep trouble, slurred speech



Patient #2: MRI C & D Symptoms: balance problems, vomiting, normal cognition



a) Label the following features in the copies of the MRIs in your handout.

1. Cerebrospinal fluid (CSF) in ventricles
2. Cerebrospinal fluid (CSF) surrounding the brain)
3. Brain tissue
4. Skull
5. The abnormal region.

b) Describe which area of the brain is affected in these individuals (General terms are ok: cortex, cerebrum, midbrain, cerebellum, brainstem, etc.).

Based on the patient symptoms, what can you deduce about the function of each area?

Patient 1: Cortex and Midbrain with midline involvement– The cortex is involved in higher-order cognition, memory, speech, and sleep.

Patient 2: Cerebellum –The cerebellum is important for balance and coordination.

c) Are the MRIs T1 or T2? What characteristics let you know?

All are T1. In general, T1- and T2- weighted images can be easily differentiated by looking at the CSF in the brain ventricles. CSF is dark on T1 and bright on T2.

Bonus: What is the difference between the two images shown for patient 1 (A & B)? How can you tell?

Image B is T1 and FLAIR: T1 with Fluid-attenuated inversion recovery (FLAIR). The edema (abnormal fluid build-up) is all bright.

d) Label the sectioning plane for each MRI (use the words “coronal”, “sagittal”, or “horizontal”)

Patient 1: A: Horizontal; B: Horizontal

Patient 2: C: Horizontal (L); D: Coronal (R).

e) Observe the specimens at each station.

Table 1: Infiltrative Astrocytoma/ Glioblastoma

Table 2: Medulloblastoma

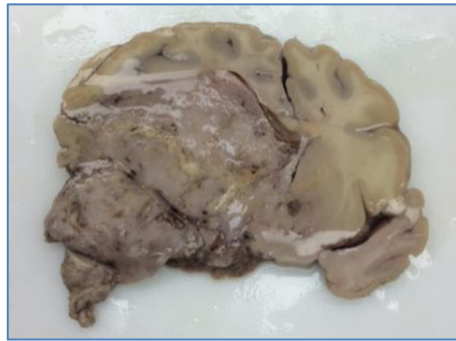
Look: Please be respectful of the patients/ families who donated these brains for teaching purposes. Please do not take pictures. Tumor tissue sometimes recruits extra blood vessels to feed the tumor. Do you see areas of more or larger blood vessels (brownish)?

Feel: If you are wearing gloves, you may touch the slices and gently press down to detect any differences in consistency. Please do not pick up the pieces. Is the tumor tissue softer than normal brain tissue?

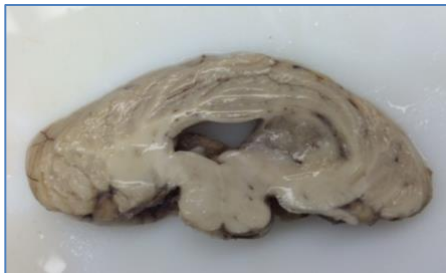
Briefly sketch out each specimen and highlight or circle the area of the brain that seems atypical.

f) Which MRI and type of tumor likely matches each patient?

Patient 1: [Table 1: Infiltrative Astrocytoma/ Glioblastoma](#)



Patient 2: [Table 2: Medulloblastoma](#)



Station 2: Bioinformatics: Changing lists to meaning

Researchers in Maria Lehtinen's lab wanted to understand the role of MYC in closure of the neural tube. They used genetic tools to overexpress MYC in the mouse brain and observed pediatric tumors. Using a technique called RNA-seq, they are able to understand how MYC overexpression causes these tumors.

Bioinformatics gives researchers tools to explore these large datasets.

When comparing the tumor brain to the normal brain, the researchers found a set of mRNAs that were higher in the brain with a tumor. These genes are listed in table 1 (tumor_brain_upreg_genes.txt).

DAVID is a website that helps researchers interpret lists of genes that are the output of common bioinformatics analyses. DAVID takes the genes as input and classifies them into general functional categories. Follow instructions below to use DAVID to investigate the role of MYC.

1. Visit the DAVID website: <https://david-d.ncicrf.gov/summary.jsp>
2. On the left-hand side, there is a blue box with 3 tabs: "Upload", "List", and "Background". Click the "Upload" tab.
 - a. It should look like this:

The screenshot shows the DAVID Functional Annotation Tool interface. The 'Upload' tab is selected. The interface is divided into a left sidebar and a main content area. The sidebar contains the following steps:

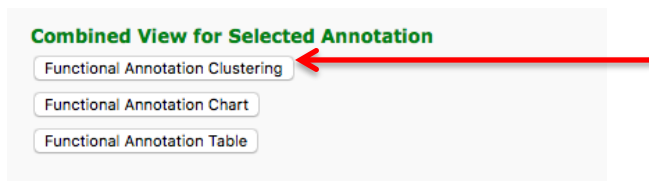
- Step 1: Enter Gene List
A: Paste a list (with a text input field and a 'Clear' button)
- Or
- B: Choose From a File
(with 'Choose File' and 'Multi-List File' buttons)
- Step 2: Select Identifier
(with a dropdown menu showing 'AFFYMETRIX_SPMFMT_ID')
- Step 3: List Type
(with radio buttons for 'Gene List' and 'Background')
- Step 4: Submit List
(with a 'Submit List' button)

The main content area contains the following sections:

- Navigation: Home, Start Analysis, Shortcut to DAVID Tools, Technical Center, Downloads & APIs, Term of Service, Why DAVID?, About Us
- Warning: *** You are currently using DAVID 6.7. ***
- Key Concepts:
 - The DAVID Gene Concept: DAVID 6.7 is designed around the "DAVID Gene Concept", a graph theory evidence-based method to agglomerate species-specific gene/protein identifiers from a variety of public genomic resources including NCBI, PIR and Uniprot/Swissprot. The DAVID Gene Concept method groups tens of million of identifiers from over 65,000 species into 1.5 million unique protein/gene records. [More](#)
 - Term/Gene Co-Occurrence Probability: Ranking functional categories based on co-occurrence with sets of genes in a gene list can rapidly aid in unraveling new biological processes associated with cellular functions and pathways. DAVID 6.7 allows investigators to sort gene categories from dozens of annotation systems. Sorting can be based either the number of genes within each category or by the EASE-score. [More](#)
 - Gene Similarity Search: Any given gene is associating with a set of annotation terms. If genes share similar set of those terms, they are most likely involved in similar biological mechanisms. The algorithm tries to group those related genes based on the agreement of sharing similar annotation terms by Kappa statistics. [More](#)
 - Term Similarity Search: Typically, a biological process/term is done by a corporation of a set of genes. If two or more biological processes are done by similar set of genes, the processes might be related in the biological network somehow. This search function is to identify the related biological processes/terms by quantitatively measuring the degree of the agreement how terms share the similar participating genes. [More](#)

3. Copy the entire contents of the file "tumor_brain_upreg_genes".
4. Paste the gene list in the "Step 1: Enter gene list (box "A"). In the "Upload" Tab.
5. Under "Step 2: Select Identifier", select "OFFICIAL_GENE_SYMBOL" from the dropdown menu
6. Under "Step 3: List Type", select the radial button next to "gene list".
7. Hit the "submit list" button. It will reload the page and generate a pop up box with an error:
 - a. "BACKGROUNDS", click "close" on the error.

8. In the blue box on the left, click the “List” tab. Under “Select to limit annotations by one or more species”, click on “Mus musculus” (scientific name for lab mouse), so that only Mus musculus is highlighted. Click “Select Species”.
 - a. On the right, you should see green text that says “103 DAVID IDs”.
 - b. Now you are ready to analyze your data.
9. Look to the right. You will see a section in the bottom in green letters called “Combined view for selected annotation”. Click on “Functional Annotation Clustering”.



- a) DAVID identifies important gene categories called clusters. What types of cellular processes (look at words in the second column) are parts of your top 3 clusters?
 - Cluster 1: [ribosome](#)
 - Cluster 2: [ribosome](#)
 - Cluster 3: [ER](#)

- b) Which cellular processes are different in the 2 conditions (healthy vs. tumor)?

Sample answers: these genes regulate mitosis- so in tumors they are highly expressed. Other possibilities? These genes may help the cells hide from the immune system, mobilize to metastasize, recruit blood vessels, or provide all the cellular building blocks the cells need to survive.

- c) Remember these genes are over expressed (higher expression than normal) in the mice with tumors (MYC overexpression). What do the clusters tell you about MYC function? i.e. What do you think the normal relationship between MYC and these genes are in a healthy brain (does MYC repress or promote the cellular processes you listed above?)

Sample answers: This suggests to me that MYC may be involved in promoting translation via inducing ribosome-making. These clusters also indicate the endoplasmic reticulum is involved—this makes proteins so another piece of evidence that translation is induced by MYC. MYC promotes translation.

Change some options. Click on the + symbol next to the word “options”.

The screenshot shows the DAVID bioinformatics tool interface. A red arrow points to the '+' icon next to the 'Options' tab. The 'Options' panel is expanded, showing settings for Classification Stringency (Medium), Kappa Similarity (3), Similarity Threshold (0.50), Initial Group Membership (3), Final Group Membership (3), Multiple Linkage Threshold (0.50), EASE (1.0), and Display options (Fold Change, Bonferroni, Benjamini, FDR, LT,PH,PT). Buttons for 'Rerun using options' and 'Create Sublist' are visible.

d) The default classification stringency is “medium”. If you change it to “highest” and hit “re-run using options” and then try “lowest”; what changes? The number / type of categories? The number of elements in each category?

Both:

The number of elements in each category: Low stringency puts more elements in one category.

The type of categories: With low stringency, the more elements are less related so the groups are more mixed.

The number of categories: Lowest = 17 clusters; Highest = 32 clusters; Medium = 29 clusters.

e) For this data set, what stringency gives you the most useful information?

None are perfect. I think the “low” is the most useful because there’s only 2 ribosome related clusters, so we can see other potentially interesting categories

Optional: Visualize the same data in another way

REVIGO is another free, online software that takes the analysis from DAVID and helps scientists visualize the results.

1. Navigate to REVIGO: <http://revigo.irb.hr/>
2. Copy the entire contents of the file “for_REVIGO.txt” and paste it into the box.
3. Under “Advanced Options” → “Select a database with GO term sizes”, click the pull down menu and select “Mus musculus”.
4. Hit “Start REVIGO”.

The first plot (scatterplot) you see takes the terms generated by DAVID and plots them based on how many of those types of terms were present in the submitted list (size of circle) and how statistically significant that term is (color of circle) in relation to each other. So many circles near each other in the graph indicate that there are many closely related processes that were found to be significant.

Hover your cursor over several of the circles to see which cellular processes are upregulated in mice with the tumors.

1. Does the evidence from REVIGO support your conclusions from DAVID?

Sample answer: yes, translation is still implicated. In fact, it's the only thing that remains significantly affected by the experiment (it's the only blue circle).

2. Can you see other biological processes implicated that you could not identify from DAVID?

Sample answer: yes, there is a circle of "Wnt-signaling pathway" and "anterior-posterior patterning".

3. Are there any circles/results that doesn't seem to fit with your understanding of the experiment? Can you guess why this might be?

Sample answer: yes, these circles that have to do with anterior/posterior patterning. This could be because those genes associated with those terms have been studied mostly via developmental patterning studies, and this is the first study to associate these genes with brain tumors. GO analysis is only a helpful starting point- we can only conclude that these genes are involved in a process based on previous studies. Further studies are necessary to follow up on this term analysis.

Explore the data! Click the "Interactive graph" tab and the "Treemap" tab to see different ways of exploring the same data. Researchers sometimes directly use these sorts of plots to represent their GO analysis.

Try on your own!

Try visualizing the list of genes that are upregulated in normal brain tissue with DAVID or REVIGO. The list of genes upregulated in normal brain tissue is called "normal_brain_upreg_genes.txt"